Supplemental Material for

Targeted Memory Reactivation During Sleep Improves Next-Day Problem Solving

Study Design Considerations

A recent meta-analytic review found that incubation effects are moderated by design factors such as problem type, length of incubation period, and type of incubation task (Sio & Ormerod, 2009). Therefore, we designed the current study to surmount challenges that may have produced inconsistencies in the sleep, incubation, and problem-solving literatures. We selected puzzles likely to require restructuring, broadly construed, as is common in incubation studies. Solving these puzzles often depends on a person ignoring initially prepotent ideas and integrating more distantly related ideas, or utilizing information that was initially ignored. Some extant lab studies presented a single problem, which could lack both sensitivity and generalizability. Other studies briefly presented many short problems of a single type, which could lead participants to confuse problem elements and, therefore, decrease the usefulness of incubation. To avoid both problems, for each participant, we collected data from 12 dissimilar puzzles they were unable to solve on the first attempt and randomly paired each puzzle with a unique sound. Additionally, the data collection was spread over two days (two evenings and two mornings), so that on any given night participants were not attempting too many puzzles or hearing too many sounds, which could lead to participants' confusion of the sound-puzzle pairings and dilute potential effects. Furthermore, the impact of sleep processing on problem solving may take time and multiple cycles of both slow-wave sleep (SWS) and rapid eye movement (REM) sleep to develop. Whereas some sleep studies with Targeted Memory Reactivation (TMR) have utilized a short nap, we presented sound cues over a full night of sleep. Finally, to increase power and control of participant-related and situational variables, we tested a relatively large number of participants in a within-participant design.

Sleep Monitoring and Cuing System Validation

The Sleep-Monitoring and Cuing System (SMCS) in our study employed an automated wireless sleep-stage monitoring device (WS, Zeo Inc). We modified the WS to use wet electrodes (silver/silver-chloride electrodes filled with electrolyte gel), which generally should improve signal quality compared to the dry electrodes used in validation studies. We targeted SWS for two reasons. First, this strategy follows the TMR literature, which largely presents cues during SWS. Second, we wanted to present cues without waking participants, which is less-likely in SWS. However, our hypotheses and outcomes are not critically tied to SWS, as being able to produce a cuing effect by presenting puzzle-associated cues in any sleep stage would be meaningful. Nonetheless, the device we utilized has a good record for targeting SWS.

In prior validation studies of the WS, agreement between the WS and human scorers on epochs of sleep versus wakefulness was 93.6%, compared to 95.8% between two human raters (Shambroom, Fábregas, & Johnstone, 2012) although compared to another automated scoring system the WS showed only 80.9% agreement (Griessenberger, Heib, Kunz, Hoedlmoser, & Schabus, 2013). In addition, the WS has high sensitivity to detect sleep (97.6%) but lower specificity (56.1%), suggesting the WS underestimates wake (Tonetti et al., 2013). As mentioned in the manuscript, in our study we observed general agreement between participants' self-reported and SMCS-reported sleep (395 minutes self-reported sleep, 95% CI = [380, 410] vs. 393 minutes SMCS-recorded sleep, 95% CI = [377, 409]) and fairly few reports of awakening during the night and hearing the puzzle-associated sounds (3.5% of trials), suggesting that the device did not regularly awaken participants.

Regarding specific sleep stages, in one study the WS agreed with a consensus score from two human PSG scorers 81.1% of the time while the two scorers agreed with each other 83.2% of the time (Cohen's kappa of .70 and .74 respectively) and showed moderate to high agreement between the scorers and the WS with positive predictive values of 85.6%, 74.4%, and 69.1% for light, REM, and deep sleep respectively (Shambroom et al., 2012). Two other validation studies also showed moderate to high agreement with expert scorers with Cohen's kappas of .56, .67, and .70 in an overnight study (Tonetti et al., 2013) and .49, .28, .72 in a nap study (Cellini, McDevitt, Ricker, Rowe, & Mednick, 2015) for light, REM, and deep sleep respectively. Together, these validation studies suggest that, despite occasional errors, overall the sleep stages indicated by the device are similar to those from human PSG scorers. The SMCS, with the improvement of wet electrodes, has also been successfully used in a prior study to present sound cues during sleep and show behavioral consequences the next day (Honma et al., 2016).

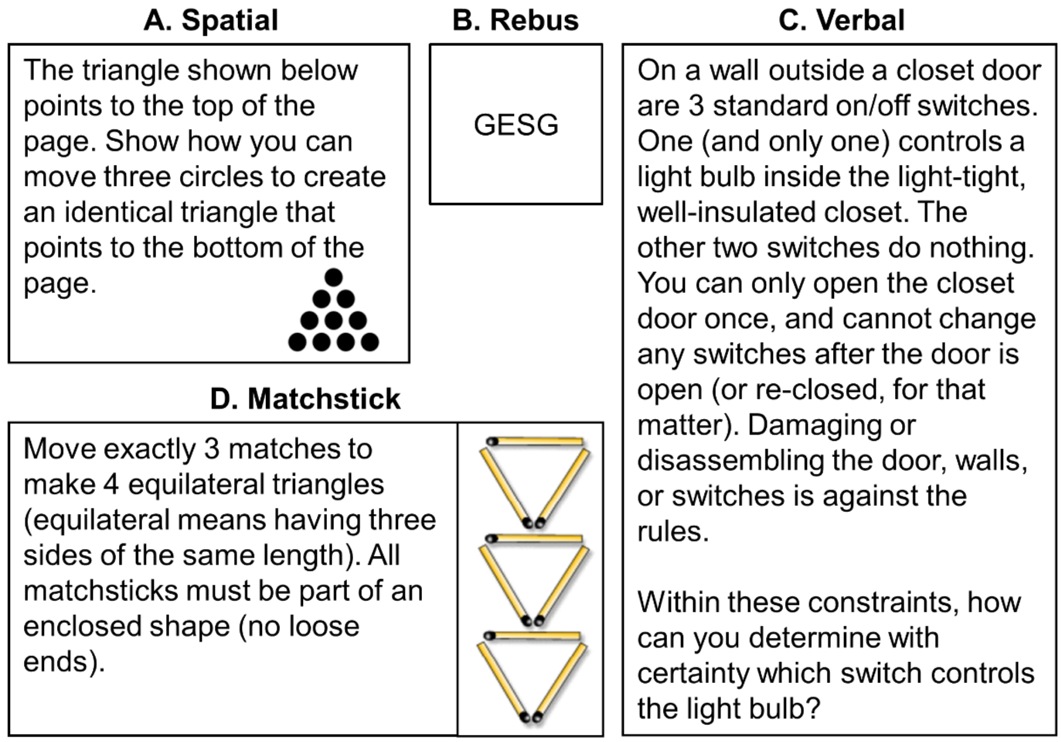

Fig. S1.

Example puzzles from each of the four puzzle types. (A) Spatial puzzles required participants to manipulate a spatial configuration to attain a goal state. (B) Rebus puzzles required participants to identify the word or phrase that the image depicted. (C) Verbal puzzles required participants to complete a specified objective based on a scenario. (D) Matchstick puzzles required participants to add, move, or remove matches to attain a configuration or goal. Answers: (A): move the top circle to the bottom center and the two bottom corner circles to be in line with the second row. (B): "scrambled eggs." (C): turn on switch 1 and leave it on for a couple minutes, then turn it off, turn on switch 2 and open the door. If the lightbulb is on, switch 2 controls it, if it is off and warm, switch 1 controls it, and if it is off and cold, switch 3 controls it. (D): move the bottom triangle so that it sits to the right of the two above triangles forming a small triangle in the space where the three matchstick triangles meet. Each triangle is equilateral, although the newly created one is smaller than the other three.

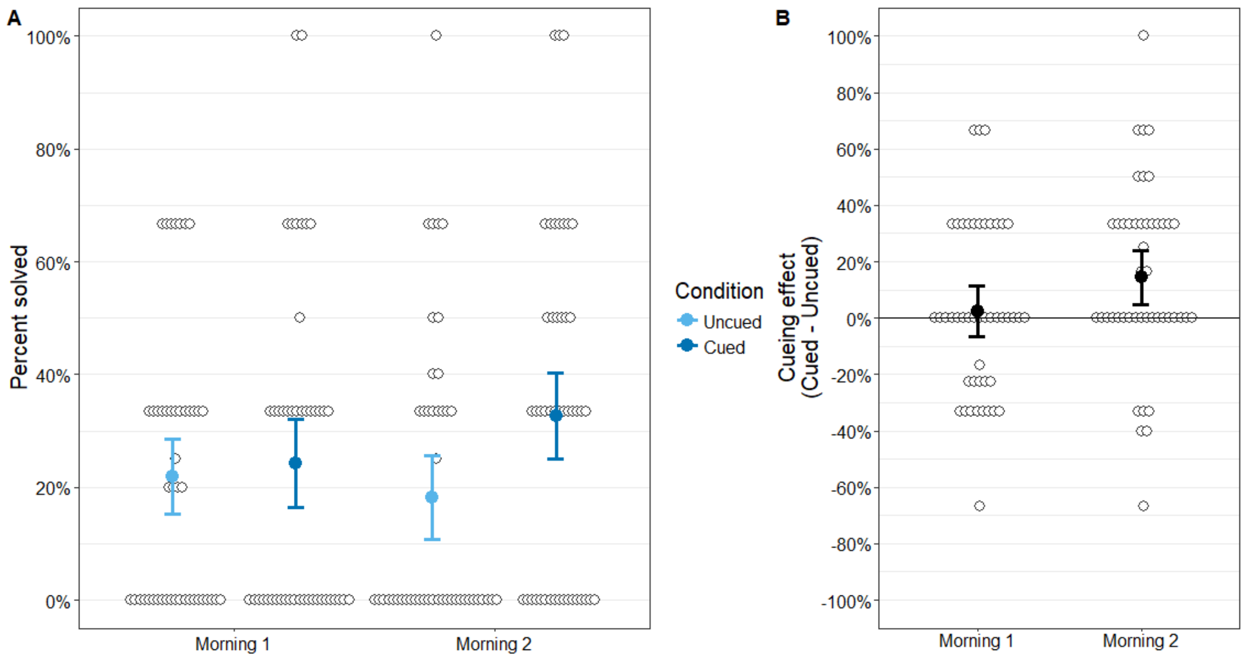

Fig. S2.

A. Solving rates for Uncued vs. Cued puzzles separately for each of the morning sessions. Error bars represent 95% within-subject confidence intervals. **B.** Within-participant cuing effect (percent Cued solving minus percent Uncued solving). Error bars represent between-subject 95% confidence intervals. A 2 (Cuing) x 2 (Session) within-subjects ANOVA shows a main effect of Cuing (F(1, 45) = 7.55, η_p^2 = .14, p = .009), no main effect of Session (F(1, 45) = 0.45, η_p^2 = .01, p = .505), and no interaction (F(1, 45) = 3.01, η_p^2 = .06, p = .090). Note that only 46 participants are included in this analysis due to missing data.

Table S1

Problem solving cuing effect by exclusion criteria

	Exclusion Criteria	Cued Mean	Cued SD	Uncued Mean	Uncued SD	t	df	p
1.	Exclude heard sounds and nights with no sounds (9.6% of data)	31.7%	28.3%	20.5%	.18.9%	3.22	56	.002
2.	None	30.0%	25.5%	20.1%	18.6%	3.29	56	.002
3.	Exclude only heard sounds (3.6% of data)	31.7%	28.3%	20.1%	18.6%	3.52	56	.001
4.	Exclude only nights with no sounds (6.0% of data)	30.0%	25.5%	20.5%	18.9%	2.97	56	.004

Note. Line 1 includes data as described in the text, excluding any data that could affect results. Line 2 includes all possible data; Lines 3 and 4 include all data except for the specific criteria noted. Regardless of exclusion criteria, participants solved more Cued than Uncued puzzles.

Table S2

Overnight sleep stage duration (minutes and percent of total sleep period) and correlations with solving and memory cuing effects

	Mean (minutes)	*SD* (minutes)	Mean (percent)	*SD* (percent)	Correlation with Solving Cuing Effect	*p*	Correlation with Memory Cuing Effect	*p*
Light Sleep	205.01	46.73	52.2%	9.3%	.05	.70	.15	.29
SWS	61.00	33.30	16.0%	9.1%	.16	.24	-.12	.37
REM	100.73	35.87	25.2%	6.6%	-.02	.87	-.18	.19
Wake	17.95	17.87	4.5%	4.2%	.04	.75	.32	.02
Missing	6.16	17.87	1.5%	3.9%	.11	.43	.13	.33
Total Sleep	393.20	59.37	----	----	.17	.22	.07	.63

Note. Correlation values are for minutes in the sleep stage and the respective cuing effects. *P* values are uncorrected for multiple comparisons. Sound cues were presented during all instances of recorded SWS.

Table S3

Problem solving cuing effect by puzzle type (matchstick, rebus, spatial, and verbal).

	N	Uncued Percent Solved	Uncued *SD*	Cued Percent Solved	Cued *SD*	Cued - Uncued Difference	*t*	*p*
Matchstick	8	28.1	21.9	39.4	19.6	11.3	2.17	0.07
Rebus	11	25.2	18.5	36.7	27.4	11.5	1.91	0.09
Spatial	11	23.4	27.9	40.1	22.5	16.7	2.35	0.04
Verbal	11	18.3	14.0	23.1	22.4	4.8	0.59	0.57

Note. *P* values are uncorrected for multiple comparisons. One matchstick puzzle is excluded from analysis because it was often solved in the evening (so it was excluded from morning datasets) and the few times it was presented in the morning, it was excluded for other reasons (e.g., the sound was heard during the night).

References

Cellini, N., McDevitt, E. A., Ricker, A. A., Rowe, K. M., & Mednick, S. C. (2015). Validation of an automated wireless system for sleep monitoring during daytime naps. *Behavioral Sleep Medicine*, *13*(2), 157–168. https://doi.org/10.1080/15402002.2013.845782

Griessenberger, H., Heib, D. P. J., Kunz, A. B., Hoedlmoser, K., & Schabus, M. (2013). Assessment of a wireless headband for automatic sleep scoring. *Sleep and Breathing*, *17*(2), 747–752. https://doi.org/10.1007/s11325-012-0757-4

Honma, M., Plass, J., Brang, D., Florczak, S. M., Grabowecky, M., & Paller, K. A. (2016). Sleeping on the rubber-hand illusion: memory reactivation during sleep facilitates multisensory recalibration. *Neuroscience of Consciousness*, *2016*(1), niw020. https://doi.org/10.1093/nc/niw020

Shambroom, J. R., Fábregas, S. E., & Johnstone, J. (2012). Validation of an automated wireless system to monitor sleep in healthy adults. *Journal of Sleep Research*, *21*(2), 221–230. https://doi.org/10.1111/j.1365-2869.2011.00944.x

Sio, U. N., & Ormerod, T. C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin*, *135*(1), 94–120. https://doi.org/10.1037/a0014212

Tonetti, L., Cellini, N., de Zambotti, M., Fabbri, M., Martoni, M., Fábregas, S. E., … Natale, V. (2013). Polysomnographic validation of a wireless dry headband technology for sleep monitoring in healthy young adults. *Physiology & Behavior*, *118*, 185–188. https://doi.org/10.1016/j.physbeh.2013.05.036

Captions for audio files

1008.mp3. Instrumental music consisting of 9.02 seconds of sound followed by 5.98 seconds of silence. Clip modified from sound obtained from freesfx.co.uk.

1036.mp3. Simple rhythm consisting of 11.15 seconds of sound followed by 3.85 seconds of silence. Clip modified from sound obtained from 1soundfx.com.